

WHAT IS CLAIMED IS:

1. A method for managing and transmitting data packets flow in a computer network system, comprising the steps of:

- (a) providing a queue having a queue level;
- 5 (b) determining the queue level;
- (c) specifying a queue hysteresis threshold and a queue low threshold;
- (d) providing a transmit probability;
- (e) if the queue level is greater than or equal to the hysteresis threshold, then dropping packets from the plurality of packets responsive to the transmit probability and
10 transmitting a remainder of the plurality of data packets to the queue;
- (f) if the queue level is less than the hysteresis threshold, then transmitting the plurality of data packets to the queue without dropping any data packets; and
- (g) if the queue level is less than the low threshold, then recalculating the transmit probability.

15 2. The method of claim 1, further comprising the steps of:

- (h) ascertaining an arrival rate of data packets into the queue;
- (i) ascertaining a sending rate of data packets out from the queue;
- (j) if the queue level is greater than or equal to the hysteresis threshold and
20 less than the sending rate from the queue, then:
 - (j)(1) decreasing the queue level until it is less than the hysteresis threshold; and

(j)(2) dropping packets from the plurality of data packets responsive to the transmit probability until the queue level decreases to at least the low threshold.

5 3. The method of claim 2, further comprising the steps of:

(k) providing a traffic flow condition parameter indicating either a first traffic flow condition or a second traffic flow condition;

(l) ascertaining the traffic flow condition parameter; and

10 wherein the step (g) of recalculating the transmit probability is performed through applying a normal transmit probability algorithm or an extended transmit probability algorithm, comprising the steps of:

(g)(1) applying the normal transmit probability algorithm to increment or decrement the transmit probability responsive to the traffic flow condition parameter indicating the first traffic flow condition; and

15 (g)(2) applying the extended transmit probability algorithm to increment or decrement the transmit probability responsive to the traffic flow condition parameter indicating the second traffic flow condition.

4. The method of claim 3, wherein the traffic flow condition parameter is a hysteresis flag having a value of ON for the first traffic flow condition and a value of OFF for the second traffic flow condition.

20

5. The method of claim 4, further comprising the step of:

if the hysteresis flag is OFF and the queue level is less than the low queue threshold and an offered load is below a link capacity, switching the hysteresis flag to ON;

5 wherein if the hysteresis flag is ON and the queue level is greater than or equal to the hysteresis threshold, then the step (e) of dropping packets and transmitting a remainder of data packets to the queue further comprises the step of switching the hysteresis flag to OFF; and

10 wherein the step (f) of transmitting the plurality of data packets to the queue without dropping any and the step (g) of recalculating the transmit probability are performed if the hysteresis flag is ON.

6. The method of claim 5 further comprising the step of providing a data flow parameter i ;

15 wherein the transmit probability is a transmit fraction T_i ; and

the step (g) of recalculating the transmit probability comprises applying a Bandwidth Allocation Transmit (BAT) algorithm, wherein:

the step (g)(1) is performed if the hysteresis flag is ON, and T_i is incremented or decremented according to the following steps:

20 (g)(1a) if $f_i(t) \leq f_{i,\min}$ then $T_i(t + dt) = \min(1, T_i(t) + w)$;
(g)(1b) else if $f_i(t) > f_{i,\max}$ then $T_i(t + dt) = T_i(t)(1-w)$;
(g)(1c) else if $B(t) = 1$ then $T_i(t + dt) = \min(1, T_i(t) + C_i B_{avg}(t))$;

(g1)(d) otherwise then $T_i(t + dt) = T_i(t)(1 - D_i O_i(t))$;

where $C_i = (S + f_{i,min} - (f_{1,min} + f_{2,min} + \dots + f_{n,min}))/16$; and

$D_i = (S - f_{i,min}) * 4$; and

the step (g)(2) if performed if the hysteresis flag is OFF, and T_i is incremented or

5 decremented according to the following steps:

(g)(2a) determining a direction of queue level;

(g)(2b) if the queue level is increasing, then $T_i = F(C_i)$; and

(g)(2c) else if the queue level is decreasing, then $T_i = G(D_i)$

where $F(C_i)$ is a BAT decreasing function and $G(D_i)$ is a BAT

10 increasing function.

7. A data flow manager structure for transmitting data packets in a computer network system, comprising:

a data flow manager in communication with a computer network;

15 a queue on the network, the queue having a queue level;

a queue hysteresis threshold;

a queue low threshold; and

a transmit probability;

wherein the data manager is configured to:

20 drop packets from a packet flow comprising a plurality of data packets responsive to the transmit probability and transmit a remainder of the flow

packets to the queue if the queue level is greater than or equal to the hysteresis threshold;

transmit the packet flow to the queue without dropping any data packets if the queue level is less than the hysteresis threshold; and

5 recalculate the transmit probability if the queue level is less than the low threshold.

8. The data flow manager structure of claim 7, wherein the data manager is further configured to decrease the queue level until it is less than the hysteresis threshold and
10 drop packets from the packet flow responsive to the transmit probability until the queue level decreases to at least the low threshold if the queue level is greater than or equal to the hysteresis threshold and less than a sending rate from the queue.

9. The data flow manager structure of claim 8, further comprising:
15 a traffic flow condition parameter indicating either a first traffic flow condition or a second traffic flow condition;

 a normal transmit probability algorithm; and

 an extended transmit probability algorithm,

 wherein the data manager is further configured to use the normal transmit
20 probability algorithm to recalculate the transmit probability responsive to the traffic flow condition parameter indicating the first traffic flow condition, and use the extended

transmit probability algorithm to recalculate the transmit probability responsive to the traffic flow condition parameter indicating the second traffic flow condition.

10. The data flow manager structure of claim 9, wherein the traffic flow condition parameter is a hysteresis flag having a value of ON for the first traffic flow condition and a value of OFF for the second traffic flow condition.

11. The data flow manager structure of claim 10, wherein the data manager is configured to:

switch the hysteresis flag to ON if the hysteresis flag is OFF and the queue level is less than the low queue threshold and an offered load is below a link capacity;

switch the hysteresis flag to OFF if the hysteresis flag is ON and the queue level is greater than or equal to the hysteresis threshold; and

transmit the plurality of data packets to the queue without dropping any and recalculate the transmit probability if the hysteresis flag is ON.

12. The data flow manager structure of claim 11 wherein the transmit probability is a transmit fraction T_i , further comprising a data flow parameter i ; and

wherein the data manager is further configured to recalculate the transmit probability by applying a Bandwidth Allocation Transmit (BAT) algorithm, wherein:

the normal transmit probability algorithm comprises:

if $f_i(t) \leq f_{i,min}$ then $T_i(t + dt) = \min(1, T_i(t) + w)$;

else if $f_i(t) > f_{i,max}$ then $T_i(t + dt) = T_i(t)(1-w)$;

else if $B(t) = 1$ then $T_i(t + dt) = \min(1, T_i(t) + C_i B_{avg}(t))$;

otherwise then $T_i(t + dt) = T_i(t)(1 - D_i O_i(t))$;

where $C_i = (S + f_{i,min} - (f_{1,min} + f_{2,min} + \dots + f_{n,min}))/16$; and

5 $D_i = (S - f_{i,min})/4$; and

the extended transmit probability algorithm comprises:

$T_i = F(C_i)$ if the queue level is increasing; and

$T_i = G(D_i)$ if the queue level is decreasing, where $F(C_i)$ is a BAT decreasing function and $G(D_i)$ is a BAT increasing function.

10

13. An article of manufacture comprising a computer usable medium having a computer readable program embodied in said medium, wherein the computer readable program, when executed on a computer, causes the computer to manage network data flow by:

15 if a queue level is greater than or equal to a hysteresis threshold, then dropping packets from flow of a plurality of packets responsive to a transmit probability and transmitting a remainder of the plurality of data packets to the queue;

if the queue level is less than the hysteresis threshold, then transmitting the plurality of data packets to the queue without dropping any data packets; and

20 if the queue level is less than a low threshold, then recalculating the transmit probability.

14. The article of manufacture of claim 13, wherein the computer readable program, when executed on a computer, further causes the computer to manage network data flow by:

if the queue level is greater than or equal to the hysteresis threshold and less than a sending rate from the queue, then:

decreasing the queue level until it is less than the hysteresis threshold; and

dropping packets from the plurality of data packets responsive to the

transmit probability until the queue level decreases to at least the low threshold.

15. The article of manufacture of claim 14, wherein the computer readable program, when executed on a computer, further causes the computer to manage network data flow by:

recalculating the transmit probability through applying a normal transmit probability algorithm to increment or decrement the transmit probability responsive to a traffic flow condition parameter indicating a first traffic flow condition; and

recalculating the transmit probability through applying an extended transmit probability algorithm to increment or decrement the transmit probability responsive to the traffic flow condition parameter indicating a second traffic flow condition.

16. The article of manufacture of claim 15, wherein the traffic flow condition parameter is a hysteresis flag having a value of ON for the first traffic flow condition and a value of OFF for the second traffic flow condition.

17. The article of manufacture of claim 16, wherein the computer readable program, when executed on a computer, further causes the computer to manage network data flow by:

5 switching the hysteresis flag to ON if the hysteresis flag is OFF and the queue level is less than the low queue threshold and an offered load is below a link capacity; switching the hysteresis flag to OFF if the hysteresis flag is ON and the queue level is greater than or equal to the hysteresis threshold; and recalculating the transmit probability if the hysteresis flag is ON.

10 18. The article of manufacture of claim 17, wherein the computer readable program, when executed on a computer, further causes the computer to manage network data flow by:

recalculating the transmit probability by applying a Bandwidth Allocation
15 Transmit (BAT) algorithm, wherein:

if the hysteresis flag is ON, T_i is incremented or decremented according to the following steps:

if $f_i(t) \leq f_{i,min}$ then $T_i(t + dt) = \min(1, T_i(t) + w)$;

else if $f_i(t) > f_{i,max}$ then $T_i(t + dt) = T_i(t)(1-w)$;

20 else if $B(t) = 1$ then $T_i(t + dt) = \min(1, T_i(t) + C_i B_{avg}(t))$;

otherwise then $T_i(t + dt) = T_i(t)(1 - D_i O_i(t))$; where $C_i = (S + f_{i,min} - (f_{1,min} + f_{2,min} + \dots + f_{n,min}))/16$; and $D_i = (S - f_{i,min}) * 4$; and

if the hysteresis flag is OFF, T_i is incremented or decremented according to the following steps:

if a queue level is increasing, then $T_i = F(C_i)$; and

else if the queue level is decreasing, then $T_i = G(D_i)$, where $F(C_i)$ is a

5 BAT decreasing function and $G(D_i)$ is a BAT increasing function;

wherein i is a data flow parameter and the transmit probability is a transmit fraction T .